1. Which of the following statements is incorrect?

Accuracy: 90/99

A. Python's str, tuple and list types are all sequence types
B. The Lists is a two-dimensional vector of elements, with a sequential relationship between the elements, accessed through the sequence numbers
C. A class can contain attributes and methods.
D. Functions allow you to write code once that can then be run whenever you need to accomplish the same task.
Correct Answer: B

2. The following ( ) types are not available for slicing operations.

Accuracy: 82/99

A, str
B. Tuple
C. dict
D. list
Correct Answer: C

3. The dictionary object’s ( ) returns a list-like view of the dictionary's values?

Accuracy: 48/99

A, items ()
B. Key ()
C. keys ()
D. values ()
Correct Answer: D

Dictionary is a built-in data type that stores data in **key-value pairs**. Each key is unique within a dictionary and is used to access its corresponding value.
e.g.

```
# Creating a dictionary
my_dict = {"name": "Alice", "age": 25, "city": "New York"}

# Using the values() method
values = my_dict.values()
```

Accuracy: 52/99

4. Which of the following statements about Python lists is correct?
A. Lists can only store integer types.
B. All elements in a list must be of the same type.
C. Elements in a list can be accessed by their index.
D. Lists are immutable data types.
Correct Answer: C

Accuracy: 70/99

5. Which of the following description of Python's loop structure is incorrect?
A. while is used to repeat a block of code based on a condition until the condition is no longer true.
B. continue only ends the current loop.
C. break is used to end the current statement, but does not jump out of the current loop.
D. for loops are typically used to iterate over any sequence type.
Correct Answer: C

6. Can logistic regression models solve linearly separable problems?
A. Yes
B. No
C. depending on the specifics of the data
D. None of the above is true
Correct Answer: A

Accuracy: 88/99

The decision boundary for logistic regression is determined by the set of points where the model outputs a **probability of 0.5**. In a linearly separable problem, logistic regression seeks **a weight vector *w*** such that this **linear equation divides the input space in a way that separates the two classes perfectly**.

7. Which of the following description of linear regression is incorrect?
A. The two variables in a correlation do not necessarily imply causation.
B. Scatter plots can visually reflect the degree of correlation between data.
C. The regression line best represents the relationship between two variables that are linearly correlated.
D. Every set of data has a regression line equation.
Correct Answer: All options are correct

8. The core of linear regression is ( )
A. modeling
B. distance measures
C. parameter learning
D. feature extraction
Correct Answer: C

Accuracy: 69/99

9. How does the Perceptron model determine the decision boundary?
A. By calculating the weighted average of the input features.
B. By setting a fixed threshold.
C. By learning to determine a linear boundary that maximizes classification accuracy.
D. By randomly selecting input features as the classification boundary.
Correct Answer: C

Accuracy: 90/99

10. In a perceptron, what happens if the data are not linearly separable?
A. The perceptron will still converge to a solution.
B. The perceptron algorithm will not converge.
C. The perceptron automatically switches to a non-linear model.
D. The perceptron will discard non-separable points.
Correct Answer: A

It will still converge, but the loss value is relatively large.

11.Judging right from wrong: Logistic regression is a supervised machine learning algorithm? Accuracy: 95/99
A) Yes
B) No
Correct Answer: A

12. Which statement is true about the logistic function used in logistic regression? Accuracy: 90/99
A. It outputs values from -1 to 1.
B. It is used to maximize the margin between classes.
C. It outputs values from 0 to 1, suitable for binary classification.
D. It provides a probability that is used to select among multiple classes.
Correct Answer: C

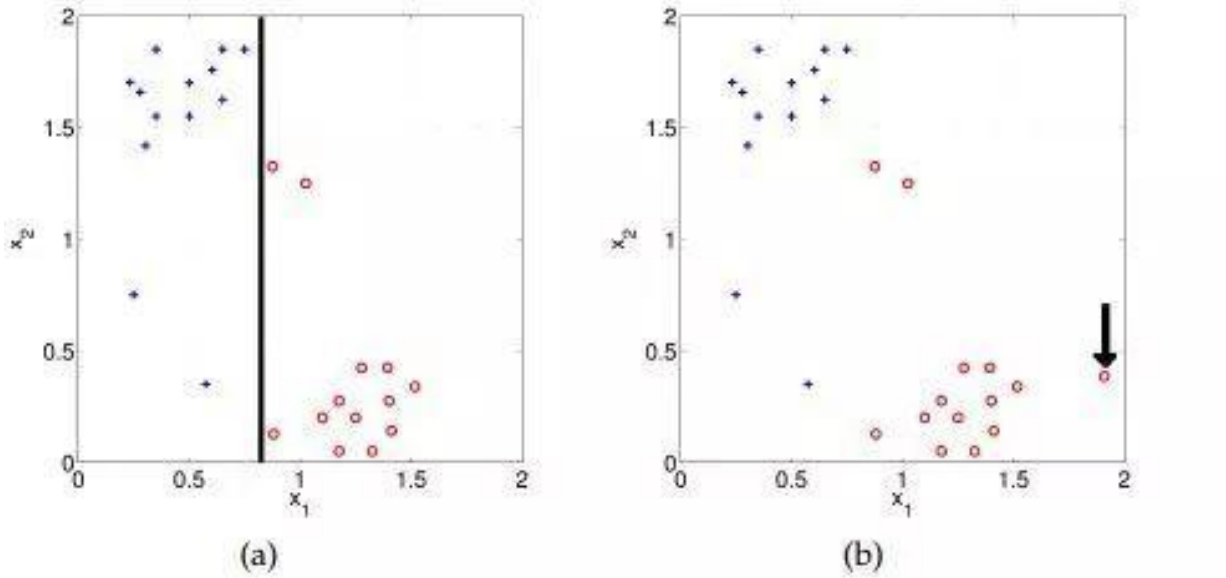
Consider the following logistic regression model: $P(y=1|x,w) = g(w0 + w1x)$, with $g(z)$ being the logistic function.
In the above equation $P(y=1|x;w)$ is viewed as a function of x, which we can get by changing the parameter w.

13. What is the range of p in this case? Accuracy: 79/99
A. (0,inf)
B. (-inf,0)
C. (0,1)
D. (-inf,inf)
Correct Answer: C
The value of x is in the real number range from $-\infty$ to $+\infty$ and the logistic function will give an output between (0,1).

14. In the above problem, which function do you think will make p between (0,1)? Accuracy: 83/99
A. Logistic functions
B. log-likelihood function
C. a mixture of both
D. neither
Correct Answer:A

15. Suppose you have given two classes "a" and "b" in two scatterplots (positive classes in blue and negative classes in red). In scatterplot a, you have correctly categorized all the data points using logistic regression (the black line is the decision boundary).

A. the bias will be high
B. the bias will be low
C. it's hard to say
D. none of the above
Correct Answer: A



16. What is the primary function of a Multi-layer Perceptron (MLP) as described in the PDF? Accuracy: 46/99

A. To linearly classify data points.
B. To approximate any function using its layered structure.
C. To reduce the dimensionality of input data.
D. To calculate linear regressions.
Correct Answer: B

17. What does an MLP with one hidden layer theoretically capable of representing? Accuracy: 40/99

A. Only linear relationships
B. Any continuous function, acting as a universal approximator.
C. Only binary classification functions.
D. Functions that require at least three layers to compute.
Correct Answer: B

18. In an MLP, what is the role of the activation function?

Accuracy: 24/99

A. To linearly combine inputs from the previous layer.
B. To add bias to each neuron's output.
C. To introduce non-linearity into the model.
D. To reduce the dimensionality of the input data.
Correct Answer: C

The primary reasons for using an activation function are:

Non-Linearity: Without a non-linear activation function, no matter how many layers the network has, it would still behave like a single-layer perceptron because the composition of **linear functions is itself a linear function.** The non-linearity introduced by activation functions allows MLPs to **learn more complex patterns in the data**.

19.What is the form of the loss function for linear regression?

Accuracy: 85/99

$$L(\theta) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

20. What is differences between linear regression and logistic regression？

Accuracy: 88/99

- Outcome Variable:
  Linear Regression: Used for predicting **a continuous outcome** variable (dependent variable).
  Logistic Regression: Used for predicting **a categorical outcome variable.**
- Loss Function:
  Linear Regression: Typically uses **Mean Squared Error** (MSE).
  Logistic Regression: **binary cross-entropy**.
- Assumptions:
  Linear Regression: The relationship between the independent and dependent variables is linear. Observations are independent of each other. The residuals (errors) of the model are normally distributed.
  Logistic Regression: The dependent variable is binary. The logit transformation (log-odds) of the probability is a linear combination of the predictors.